

APPLYING METRIC REGULARITY TO COMPUTE A CONDITION MEASURE OF A SMOOTHING ALGORITHM FOR MATRIX GAMES

BORIS S. MORDUKHOVICH¹, JAVIER F. PEÑA² and VERA ROSHCHINA³

Abstract. We develop an approach of variational analysis and generalized differentiation to conditioning issues for two-person zero-sum matrix games. Our major results establish precise relationships between a certain condition measure of the smoothing first-order algorithm proposed by Gilpin et al. [*Proceedings of the 23rd AAAI Conference (2008) pp. 75–82*] and the exact bound of metric regularity for an associated set-valued mapping. In this way we compute the aforementioned condition measure in terms of the initial matrix game data.

Key words. matrix games, smoothing algorithm, condition measure, variational analysis, metric regularity, generalized differentiation

AMS subject classifications. 90D10, 90C30, 49J53, 49J52, 49M45

Abbreviated title. Condition measure for matrix games

1. Introduction and formulation of main results. This paper is devoted to applications of advanced techniques in variational analysis and generalized differentiation to the study of conditioning in optimization. Our specific goal is to apply the key notions and generalized differential characterizations of Lipschitzian stability and metric regularity, fundamental in variational analysis, to computing a certain condition measure of the first-order smoothing algorithm proposed in [4] to find approximate Nash equilibria of two-person zero-sum matrix games.

To the best of our knowledge, applications of Lipschitzian stability and metric regularity to numerical aspects of optimization were initiated by Robinson in the 1970s; see, e.g., [17] and the references therein. In the complexity theory, Renegar [15, 16] established relationships between the rate of convergence of interior-point methods for linear and conic convex programs and their “distance to ill-posedness” and related condition numbers. We refer the reader to [1, 2, 3, 6, 7, 14] and their bibliographies for more recent results in this direction for various algorithms in convex and nonconvex optimization problems.

In [4], a new condition measure was introduced to evaluate the complexity of a first-order algorithm for solving a two-person zero-sum game

$$\min_{x \in Q_1} \max_{y \in Q_2} x^T A y = \max_{y \in Q_2} \min_{x \in Q_1} x^T A y, \quad (1.1)$$

where $A \in \mathbb{R}^{m \times n}$, where the symbol T stands for transposition, and where each of the sets Q_1 and Q_2 is either a simplex (in the matrix game formulation) or a more elaborate polytope (in the case of sequential games). Problems of this type arise in many interesting applications; see, e.g., [13, 20, 22, 23] and the references therein.

It was shown in [4] that an iterative version of Nesterov’s first-order smoothing algorithm [11, 12] computes an ε -equilibrium point (in the sense of Nash) for problem (1.1) in $\mathcal{O}(\|A\| \kappa(A) \ln(1/\varepsilon))$ iterations, where $\kappa(A)$ is a condition measure of (1.1) depending only on A ; see (1.7) for the definition of the condition measure $\kappa(A)$ in the case of matrix games. The dependence of this complexity bound on ε is exponentially better than the complexity bound $\mathcal{O}(1/\varepsilon)$ in the original Nesterov’s smoothing techniques. Furthermore, it was proved in [4] that the condition measure

¹Wayne State University, Detroit, MI, USA (boris@math.wayne.edu). Research of this author was partly supported by the US National Science Foundation under grants DMS-0603846 and DMS-1007132 and by the Australian Research Council under grant DP-12092508.

²Carnegie Mellon University, Pittsburgh, PA, USA (jfp@andrew.cmu.edu). Research of this author was partly supported by the US National Science Foundation under grant CCF-0830533.

³Universidade de Évora, Évora, Portugal (veraroshchina@gmail.com).

$\kappa(A)$ is always finite while the proof therein was non-constructive. In particular, no explicit upper bound on $\kappa(A)$ was given. However, numerical results reported in [4] clearly demonstrate that the developed iterative version of Nesterov's smoothing algorithm is faster than other algorithms known in this setting and that this version exhibits at least linear convergence for a random collection of the problem instances considered in [4]. This allows us to treat the number $\kappa(A)$ as a condition measure of the algorithm and evaluate it in what follows.

In this paper we focus on the *matrix game equilibrium problem*

$$\min_{x \in \Delta_m} \max_{y \in \Delta_n} x^T A y = \max_{y \in \Delta_n} \min_{x \in \Delta_m} x^T A y, \quad (1.2)$$

where the m -dimensional simplex

$$\Delta_m := \left\{ x \in \mathbb{R}^m \mid \sum_{i=1}^m x_i = 1, x \geq 0 \right\}$$

describes the set of mixed strategies for the x -player (Player 1) with m pure strategies; similarly for the y -player $y \in \Delta_n$ (Player 2). This means that if Player 1 uses $x \in \Delta_m$ and Player 2 uses $y \in \Delta_n$, then Player 1 gets payoff $-x^T A y$ while Player 2 gets payoff $x^T A y$. Thus the equilibrium problem (1.2) can be reformulated as the following problem of *nonsmooth convex optimization*:

$$\text{minimize } F(x, y) \text{ subject to } (x, y) \in \Delta_m \times \Delta_n, \quad (1.3)$$

where the minimizing cost function $F(x, y)$ is defined by the maximum

$$F(x, y) := \max \{ x^T A v - u^T A y \mid (u, v) \in \Delta_m \times \Delta_n \}. \quad (1.4)$$

It is easy to observe that

$$\min \{ F(x, y) \mid (x, y) \in \Delta_m \times \Delta_n \} = 0. \quad (1.5)$$

Taking (1.5) into account, we say [4, 23] that a feasible pair $(\bar{x}, \bar{y}) \in \Delta_m \times \Delta_n$ is a *Nash equilibrium* to (1.2) if $F(\bar{x}, \bar{y}) = 0$, which corresponds to an optimal solution of the constrained optimization problem (1.3). Consider the optimal solution set

$$S := \{ (\bar{x}, \bar{y}) \in \Delta_m \times \Delta_n \mid F(\bar{x}, \bar{y}) = 0 \} = F^{-1}(0) \cap (\Delta_m \times \Delta_n) \quad (1.6)$$

and, following [4], define the *condition measure* $\kappa(A)$ of the matrix game (1.2) depending on the underlying matrix A via the objective (1.4) and the optimal solution set (1.6) as

$$\kappa(A) := \inf \{ \kappa \geq 0 \mid \text{dist}((x, y); S) \leq \kappa F(x, y) \text{ for all } (x, y) \in \Delta_m \times \Delta_n \}, \quad (1.7)$$

where $\text{dist}(\cdot; S)$ stands for the standard Euclidean distance function.

In what follows we derive three major results concerning the characterization of the condition measure $\kappa(A)$ in (1.7). The first theorem shows that the condition measure $\kappa(A)$ precisely relates to the exact bound of metric regularity for an associated set-valued mapping built upon the cost function (1.4). The second result expresses this exact regularity bound via the subdifferential of the convex function (1.4) and the normal cone to the simplex product $\Delta_m \times \Delta_n$ and then computes the latter constructions in terms of the initial data of (1.2). Finally, we arrive at an exact formula for evaluating $\kappa(A)$, which is a key step towards performing further complexity analysis of the algorithm [4].

To formulate the first theorem, define a set-valued mapping $\Phi: \mathbb{R}^{m+n} \rightrightarrows \mathbb{R}$ by

$$\Phi(x, y) := \begin{cases} [F(x, y), \infty) & \text{if } (x, y) \in \Delta_m \times \Delta_n, \\ \emptyset & \text{otherwise} \end{cases} \quad (1.8)$$

via the cost function F constructed in (1.4). Let $\text{reg } \Phi((x, y), F(x, y))$ be the *exact bound of metric regularity* (or the *exact regularity bound/modulus*) of the mapping Φ around the point $((x, y), F(x, y)) \in \text{gph } \Phi$; see [10, 19]. For the reader's convenience we recall these concepts in Section 2 below.

THEOREM 1.1. (condition measure via the exact regularity bound). *Assume that $(\Delta_m \times \Delta_n) \setminus S \neq \emptyset$ with S defined in (1.6). Then we have the precise relationship*

$$\kappa(A) = \sup_{(x, y) \in (\Delta_m \times \Delta_n) \setminus S} \text{reg } \Phi((x, y), F(x, y)) \quad (1.9)$$

between the condition measure (1.7) and the exact regularity bound of (1.8).

Our second major result gives a complete characterization of metric regularity for the set-valued mapping Φ defined in (1.8).

THEOREM 1.2. (computing the exact bound of metric regularity). *For any point $(x, y) \in (\Delta_m \times \Delta_n) \setminus S$, the exact regularity bound of the mapping Φ from (1.8) around the point $((x, y), F(x, y))$ admits the representation*

$$\text{reg } \Phi((x, y), F(x, y)) = \frac{1}{\text{dist}(0; \partial F(x, y) + N_{\Delta_m \times \Delta_n}(x, y))} \quad (1.10)$$

via the subdifferential of the convex function (1.4) and the normal cone to the simplex product $\Delta_m \times \Delta_n$ at (x, y) .

Unifying the results of Theorem 1.1 and Theorem 1.2 and then explicitly computing the subdifferential and normal cone on the right-hand side of (1.10), we get the precise formula for computing the condition measure of the smoothing algorithm for matrix games as stated in Theorem 1.3 below. To formulate this third major result, we introduce some convenient notation. Let a_i as $i = 1, \dots, n$ and $-b_k^T$ as $k = 1, \dots, m$ stand for the columns and the rows of the matrix A , respectively. By e_j , $j = 1, \dots, m + n$, we denote the unit vectors in \mathbb{R}^{m+n} , i.e.,

$$(e_j)_l = 0 \text{ for all } l \neq j \text{ and } (e_j)_j = 1 \text{ as } j = 1, \dots, m + n.$$

For a positive integer p , let $\mathbf{1}_p := [1 \ \dots \ 1] \in \mathbb{R}^p$. Finally, given a feasible point $(x, y) \in \Delta_m \times \Delta_n$, define the index sets $I(x)$, $K(y)$, and $J(x, y)$ by

$$\begin{cases} I(x) := \left\{ \bar{i} \in \{1, \dots, n\} \mid a_{\bar{i}}^T x = \max_{i \in \{1, \dots, n\}} a_i^T x \right\}, \\ K(y) := \left\{ \bar{k} \in \{1, \dots, m\} \mid b_{\bar{k}}^T y = \max_{k \in \{1, \dots, m\}} b_k^T y \right\}, \\ J(x, y) := \{j \in \{1, \dots, m\} \mid x_j = 0\} \cup \{j = m + p \mid y_p = 0\}. \end{cases} \quad (1.11)$$

THEOREM 1.3. (computing the condition measure). *Let $(\Delta_m \times \Delta_n) \setminus S \neq \emptyset$. Then, in the notation above, the condition measure $\kappa(A)$ defined in (1.7) is computed by*

$$\kappa(A) = \sup_{(x, y) \in (\Delta_m \times \Delta_n) \setminus S} \left[\text{dist} \left(0; \text{co} \{ (a_i, b_k) \mid i \in I(x), k \in K(y) \} \right. \right. \\ \left. \left. + \text{span} \{ \mathbf{1}_m \} \times \text{span} \{ \mathbf{1}_n \} - \text{cone} [\text{co} \{ e_j \mid j \in J(x, y) \}] \right) \right]^{-1},$$

where the symbols span , cone , and co stand respectively for the linear, conic, and convex hulls of the sets in question.

The proofs of Theorem 1.1 and Theorem 1.2 given below are based on applying advanced techniques of variational analysis and generalized differentiation. This approach leads us therefore to deriving the precise formula for the condition measure in Theorem 1.3. For additional insight, we also present a direct, independent proof of the latter theorem relying on more conventional while

somewhat more laborious techniques of convex optimization employing particularly Lagrangian duality.

REMARK 1.4. (numerical implementation and further research). Numerical implementation of the formula for the condition measure in Theorem 1.3 is not a purpose of this paper and in fact is not an easy job. It has been well recognized in complexity theory that evaluating condition measures may be in general as difficult as to solve the original problem. This is true, e.g., in the cases of such fundamental complexity measures as the condition number of a matrix [5] used in estimating complexity of numerical linear algebra algorithms, Renegar’s condition number [14, 15, 16] that characterizes difficulty of solving conic feasibility problems, the “measure of condition” for finding zeros of complex polynomials introduced by Shub and Smale [21], etc.

The main purpose of this paper is not obtaining an easily computable expression for the condition measure $\kappa(A)$, but rather gaining a better understanding on how exactly the problem data influence the condition measure. Observe that the formula for $\kappa(A)$ obtained in Theorem 1.3 is much easier to evaluate and analyze than the original construction (1.7). This is valuable for the average-case and smoothed analysis of the algorithm, singling out classes of well-conditioned problems, preconditioning issues, and making further improvements to the algorithm. We will pursue these goals in our subsequent research.

The rest of the paper is organized as follows. In Section 2 we recall some basic definitions and facts of variational analysis and generalized differentiation crucial for deriving the main results of the paper. Section 3 is devoted to variational proofs of the main results formulated above. Finally, in Section 4 we present an alternative direct proof of Theorem 1.3 by employing tools of convex optimization.

Throughout the paper we use standard notation and terminology of variational analysis; see, e.g., the basic texts [10, 19].

2. Preliminaries from variational analysis and generalized differentiation. Here we confine ourselves to finite-dimensional Euclidean spaces sufficient for the subsequent applications. The reader is referred to [10, 19] for more details and related material.

Given a set-valued mapping $G: \mathbb{R}^n \rightrightarrows \mathbb{R}^m$, consider its inverse $G^{-1}: \mathbb{R}^m \rightrightarrows \mathbb{R}^n$ with $x \in G^{-1}(z) \iff z \in G(x)$ as well as its graphs, domain, and range defined respectively by

$$\text{gph } G := \{(x, z) \mid z \in G(x)\}, \quad \text{dom } G := \{x \mid G(x) \neq \emptyset\}, \quad \text{rge } G := \text{dom } G^{-1}.$$

The notion of metric regularity is of primary interest in our development.

DEFINITION 2.1. (metric regularity). A set-valued mapping $G: \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ is METRICALLY REGULAR around $(\bar{x}, \bar{z}) \in \text{gph } G$ with modulus $\mu \geq 0$ if there exist neighborhoods U of \bar{x} and V of \bar{z} such that

$$\text{dist}(x; G^{-1}(z)) \leq \mu \text{dist}(z; G(x)) \quad \text{whenever } x \in U \text{ and } z \in V. \quad (2.1)$$

The infimum of $\mu \geq 0$ over all (μ, U, V) for which (2.1) holds is called the EXACT REGULARITY BOUND of G around (\bar{x}, \bar{z}) and is denoted by $\text{reg } G(\bar{x}, \bar{z})$.

It is well known in variational analysis that the fundamental property of metric regularity is closely related to Lipschitzian behavior of inverse mappings. Recall that a mapping $G: \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ is Lipschitz-like (or has the Aubin property) around $(\bar{x}, \bar{z}) \in \text{gph } G$ with modulus $\ell \geq 0$ if there are neighborhoods U of \bar{x} and V of \bar{z} such that

$$G(x) \cap V \subset G(u) + \ell \|x - u\| \mathcal{B} \quad \text{for all } x, u \in U, \quad (2.2)$$

where \mathcal{B} stands for the Euclidean closed unit ball of the space in question. The infimum of $\ell \geq 0$ over all the combinations (ℓ, U, V) for which (2.2) holds is called the exact Lipschitzian bound of G around (\bar{x}, \bar{z}) and is denoted by $\text{lip } G(\bar{x}, \bar{z})$.

The following result can be found, e.g., in [10, Theorem 1.49].

PROPOSITION 2.2. (relationships between metric regularity and Lipschitz-like properties). Let $G: \mathbb{R}^n \rightrightarrows \mathbb{R}^m$, and let $(\bar{x}, \bar{z}) \in \text{gph } G$. Then the mapping G is metrically regular

around (\bar{x}, \bar{z}) if and only if its inverse $G^{-1}: \mathbb{R}^m \rightrightarrows \mathbb{R}^n$ is Lipschitz-like around (\bar{z}, \bar{x}) . Furthermore, we have the equality

$$\text{reg } G(\bar{x}, \bar{z}) = \text{lip } G^{-1}(\bar{z}, \bar{x}).$$

One of the key advantages of modern variational analysis is the possibility to completely characterize Lipschitzian and metric regularity properties of set-valued mappings in terms of appropriate generalized differential constructions enjoying full calculus. Let us recall such constructions used in this paper.

Given a nonempty subset $\Omega \subset \mathbb{R}^n$ and a point $\bar{x} \in \Omega$, define the Fréchet/regular normal cone to Ω at \bar{x} by

$$\widehat{N}_\Omega(\bar{x}) := \left\{ v \in \mathbb{R}^n \mid \limsup_{x \xrightarrow{\Omega} \bar{x}} \frac{\langle v, x - \bar{x} \rangle}{\|x - \bar{x}\|} \leq 0 \right\},$$

where the symbol $x \xrightarrow{\Omega} \bar{x}$ means that $x \rightarrow \bar{x}$ with $x \in \Omega$. Then the Mordukhovich (basic/limiting) normal cone to Ω at $\bar{x} \in \Omega$ is defined by

$$N_\Omega(\bar{x}) := \text{Lim sup}_{x \xrightarrow{\Omega} \bar{x}} \widehat{N}_\Omega(x), \quad (2.3)$$

where ‘Lim sup’ stands for the Painlevé-Kuratowski outer/upper limit of a set-valued mapping $M: \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ given by

$$\text{Lim sup}_{x \rightarrow \bar{x}} M(x) := \left\{ v \in \mathbb{R}^m \mid \begin{array}{l} \exists x_k \rightarrow \bar{x}, v_k \rightarrow v \text{ as } k \rightarrow \infty \text{ such that} \\ v_k \in M(x_k) \text{ for all } k = \mathbb{N} := \{1, 2, \dots\} \end{array} \right\}.$$

If the set Ω is locally closed around \bar{x} , the normal cone (2.3) admits the equivalent description (which was in fact the original definition in [8])

$$N_\Omega(\bar{x}) = \text{Lim sup}_{x \rightarrow \bar{x}} [\text{cone}(x - \Pi_\Omega(x))]$$

in terms of the projection operator $\Pi_\Omega(x) := \{y \in \Omega \mid \|y - x\| = \text{dist}(x; \Omega)\}$.

Note that the normal cone (2.3) may be nonconvex even for simple sets $\Omega \subset \mathbb{R}^n$, e.g., for the graph of $|x|$ and the epigraph of $-|x|$ at $(0, 0)$. Due to its nonconvexity, the normal cone (2.3) cannot be polar to any tangent cone. Nevertheless, this nonconvex normal cone and the corresponding subdifferential/coderivative constructions for extended-real-valued (i.e., with values in $(-\infty, \infty]$) functions and set-valued mappings satisfy comprehensive calculus rules, which are derived by using variational arguments, particularly the extremal principle of variational analysis; see [10, 19] and the references therein.

A set Ω is called *normally regular* at $\bar{x} \in \Omega$ if $N_\Omega(\bar{x}) = \widehat{N}_\Omega(\bar{x})$. The class of normally regular sets covers “nice” sets having a local convex-like structure. A major example is provided by convex sets; see, e.g., [10, Proposition 1.5].

PROPOSITION 2.3. (normal regularity of convex sets). *Let $\Omega \subset \mathbb{R}^n$ be convex. Then it is normally regular at every point $\bar{x} \in \Omega$, and its normal cone (2.3) reduces to the normal cone in the sense of convex analysis:*

$$N_\Omega(\bar{x}) = \{v \in \mathbb{R}^n \mid \langle v, x - \bar{x} \rangle \leq 0 \text{ for all } x \in \Omega\}.$$

Given next a set-valued mapping $G: \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ and a point $(\bar{x}, \bar{z}) \in \text{gph } G$, define a generalized derivative of G at (\bar{x}, \bar{z}) induced by the normal cone (2.3) to $\text{gph } G$ at (\bar{x}, \bar{z}) . Namely, the *coderivative* of G at (\bar{x}, \bar{z}) is a set-valued mapping $D^*G(\bar{x}, \bar{z}): \mathbb{R}^m \rightrightarrows \mathbb{R}^n$ with the values

$$D^*G(\bar{x}, \bar{z})(v) := \{u \in \mathbb{R}^n \mid (u, -v) \in N_{\text{gph } G}(\bar{x}, \bar{z})\}. \quad (2.4)$$

Observe that $0 \in D^*G(\bar{x}, \bar{z})(0)$ and $D^*G(\bar{x}, \bar{z})(\lambda v) = \lambda D^*G(\bar{x}, \bar{z})(v)$ for every $\lambda > 0$, i.e., the coderivative (2.4) is a positively homogeneous mapping. If $G: \mathbb{R}^n \rightarrow \mathbb{R}^m$ is single-valued and smooth around \bar{x} with the derivative $\nabla G(\bar{x})$, we have (see, e.g., by [10, Theorem 1.38])

$$D^*G(\bar{x})(v) = \{\nabla G(\bar{x})^T v\} \quad \text{for all } v \in \mathbb{R}^m.$$

The latter signifies that the coderivative (2.4) is an appropriate extension of the adjoint/transpose derivative operator to the case of nonsmooth and set-valued mappings. Note also that, by the non-convexity of the normal cone (2.3), the coderivative (2.4) is not dual to any tangentially generated graphical derivative, except of the case when G is *graphical regular* at (\bar{x}, \bar{z}) meaning that

$$N_{\text{gph } G}(\bar{x}, \bar{z}) = \widehat{N}_{\text{gph } G}(\bar{x}, \bar{z}).$$

As mentioned above, the coderivative (2.4) satisfies comprehensive calculus rules for general set-valued mappings. In this paper we only need the following one, which is a consequence of [10, Proposition 3.12]. To formulate it, recall that the indicator mapping $\delta_\Omega: \mathbb{R}^n \rightarrow \mathbb{R}$ of a set $\Omega \subset \mathbb{R}^n$ is defined by

$$\delta_\Omega(x) := \begin{cases} 0 & \text{if } x \in \Omega, \\ \emptyset & \text{otherwise,} \end{cases}$$

(a bit different from the indicator functions) and that we easily have the relationship

$$D^*(\delta_\Omega)(\bar{x})(v) = N_\Omega(\bar{x}) \quad \text{for any } \bar{x} \in \Omega \text{ and } v \in \mathbb{R}.$$

PROPOSITION 2.4. (coderivative sum rule). *Let $\Omega \subset \mathbb{R}^n$ be locally closed around $\bar{x} \in \Omega$, and let $G: \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ be closed-graph and Lipschitz-like around $(\bar{x}, \bar{z}) \in \text{gph } G$. Then for all $v \in \mathbb{R}^m$ we have the inclusion*

$$D^*(G + \delta_\Omega)(\bar{x}, \bar{z})(v) \subset D^*G(\bar{x}, \bar{z})(v) + N_\Omega(\bar{x}),$$

which holds as equality if Ω is normally regular at \bar{x} and G is graphically regular at (\bar{x}, \bar{z}) .

In what follows we employ the norm of the coderivative as a positively homogeneous mapping. The norm of a positively homogeneous mapping $M: \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ is defined by

$$\|M\| := \sup \{\|u\| \mid u \in M(v) \text{ with } \|v\| \leq 1\}$$

and admits (by passing to the inverse) the useful distance function representation below established in [3, Proposition 2.5].

PROPOSITION 2.5. (norm of positively homogeneous mappings). *Let $M: \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ be positively homogeneous. Then the norm of its inverse is computed by*

$$\|M^{-1}\| = \sup_{\|v\|=1} \frac{1}{\text{dist}(0; M(v))}.$$

The final and most important result presented in this section provides a complete coderivative characterization of the Lipschitz-like property (known as the Mordukhovich criterion [19]) with

computing the exact bound of Lipschitzian moduli; see [9, Theorem 5.7], [10, Theorem 4.10], and [19, Theorem 9.40] for different proofs.

THEOREM 2.6. (coderivative characterization of the Lipschitz-like property for set-valued mappings). *Let $G: \mathbb{R}^m \rightrightarrows \mathbb{R}^n$ be closed-graph around $(\bar{x}, \bar{z}) \in \text{gph } G$. Then G is Lipschitz-like around this point if and only if $D^*G(\bar{x}, \bar{z})(0) = \{0\}$. In this case*

$$\text{lip } G(\bar{x}, \bar{z}) = \|D^*G(\bar{x}, \bar{z})\|.$$

3. Proofs of main results. We give here complete proofs of Theorem 1.1 and Theorem 1.2, and thus derive the condition measure formula of Theorem 1.3 by variational arguments.

Let us start with the proof of Theorem 1.1. To proceed, we first establish a more convenient representation of the condition measure (1.7) for our further analysis, which in turn is preceded by a technical claim.

Observe that the function $F(x, y)$ defined by (1.4) can be written as follows:

$$F(x, y) = \max_{\substack{i=1, \dots, n \\ j=1, \dots, m}} (a_i^T x + b_j^T y).$$

In addition we represent the simplex product $\Delta_m \times \Delta_n$ by:

$$\Delta_m \times \Delta_n = \{w = (x, y) \mid w \geq 0, Ew = f\} \quad \text{with} \quad E := \begin{bmatrix} \mathbf{1}_m^T & 0 \\ 0 & \mathbf{1}_n^T \end{bmatrix} \quad \text{and} \quad f := \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

To simplify notation, rewrite the function F as

$$F(w) = \max_{\ell \in L} c_\ell^T w,$$

where $L := \{1, \dots, n\} \times \{1, \dots, m\}$ and $c_\ell^T := [a_i^T \ b_j^T]$ for each $\ell = (i, j) \in L$. Denote further $\Omega := \Delta_n \times \Delta_m = \{w \mid w \geq 0, Ew = f\}$ and rewrite (1.7) as

$$k_F := \inf \{k \geq 0 \mid \text{dist}(w; S) \leq kF(w) \text{ for all } w \in \Omega\} \quad (3.1)$$

with S given by (1.6). Observe that $\min_\Omega F(x) = 0$ by (1.5). It is also convenient for us to define the moving sets

$$S(z) := \{w \in \Omega \mid F(w) = z\} = F^{-1}(z) \cap \Omega \quad \text{with thus} \quad S = S(0) \quad (3.2)$$

and to represent the mapping Φ in (1.8) and its inverse by

$$\Phi(w) = [F(w), \infty) + \delta_\Omega(w) \quad \text{and} \quad \Phi^{-1}(z) = \{w \in \Omega \mid F(w) \leq z\}. \quad (3.3)$$

Let us finally denote $\mathcal{J} := \{1, \dots, n, n+1, \dots, m+n\}$ and define the corresponding counterparts of the index sets $I(\cdot)$ and $J(\cdot)$ from (1.11) given by

$$\mathcal{I}(w) := \{\ell \in L \mid c_\ell^T w = F(w)\}, \quad \mathcal{J}(w) := \{j \in \mathcal{J} \mid w_j = 0\}.$$

It is not difficult to verify the following technical claim, where \mathcal{B}_γ stands for the closed ball of radius $\gamma > 0$ centered at the origin.

CLAIM 3.1. (relationships between index sets). *For every $\bar{w} \in \Omega$ there exists $\gamma > 0$ such that $\mathcal{I}(w) \subset \mathcal{I}(\bar{w})$ and $\mathcal{J}(w) \subset \mathcal{J}(\bar{w})$ whenever $w \in \bar{w} + \mathcal{B}_\gamma$. **Proof.** Fix an arbitrary element $\bar{w} \in \Omega$ and let*

$$0 < \gamma < \min \left\{ \min_{\substack{\ell \in L, \\ c_\ell \neq 0}} \frac{1}{2\|c_\ell\|} \min_{i \in L \setminus \mathcal{I}(\bar{w})} (F(\bar{w}) - c_i^T \bar{w}), \min_{j \in \mathcal{J} \setminus \mathcal{J}(\bar{w})} \bar{w}_j \right\},$$

where $\min \emptyset = \infty$ by convention. It is easy to observe that such a number γ always exists. When $\mathcal{I}(\bar{w}) = L$, the inclusion $\mathcal{I}(w) \subset \mathcal{I}(\bar{w})$ is obvious. Assume thus that $L \setminus \mathcal{I}(\bar{w}) \neq \emptyset$. For every $\ell_0 \in L \setminus \mathcal{I}(\bar{w})$ and every $w \in \bar{w} + \mathcal{B}_\gamma$ we have

$$c_{\ell_0}^T w = c_{\ell_0}^T \bar{w} + c_{\ell_0}^T (w - \bar{w}) \leq c_{\ell_0}^T \bar{w} + \|c_{\ell_0}\| \gamma < c_{\ell_0}^T \bar{w} + \frac{1}{2} \min_{\ell \in L \setminus \mathcal{I}(\bar{w})} (F(\bar{w}) - c_\ell^T \bar{w}),$$

which implies the relationships

$$\begin{aligned} \max_{\ell \in L \setminus \mathcal{I}(\bar{w})} (c_\ell^T w) &< \max_{\ell \in L \setminus \mathcal{I}(\bar{w})} (c_\ell^T \bar{w}) + \frac{1}{2} \min_{\ell \in L \setminus \mathcal{I}(\bar{w})} (F(\bar{w}) - c_\ell^T \bar{w}) \\ &= \frac{1}{2} \left(F(\bar{w}) + \max_{\ell \in L \setminus \mathcal{I}(\bar{w})} c_\ell^T \bar{w} \right). \end{aligned} \quad (3.4)$$

Similarly, for every $\ell_0 \in \mathcal{I}(\bar{w})$ we have

$$c_{\ell_0}^T w = c_{\ell_0}^T \bar{w} + c_{\ell_0}^T (w - \bar{w}) \geq c_{\ell_0}^T \bar{w} - \|c_{\ell_0}\| \gamma > c_{\ell_0}^T \bar{w} - \frac{1}{2} \min_{\ell \in L \setminus \mathcal{I}(\bar{w})} (F(\bar{w}) - c_\ell^T \bar{w}),$$

which in turn implies that

$$\begin{aligned} \max_{\ell \in \mathcal{I}(\bar{w})} (c_\ell^T w) &> \max_{\ell \in \mathcal{I}(\bar{w})} (c_\ell^T \bar{w}) - \frac{1}{2} \min_{\ell \in L \setminus \mathcal{I}(\bar{w})} (F(\bar{w}) - c_\ell^T \bar{w}) \\ &= \max_{\ell \in \mathcal{I}(\bar{w})} (c_\ell^T \bar{w}) - \frac{1}{2} F(\bar{w}) + \frac{1}{2} \max_{\ell \in L \setminus \mathcal{I}(\bar{w})} (c_\ell^T \bar{w}) \\ &= \frac{1}{2} \left(F(\bar{w}) + \max_{\ell \in L \setminus \mathcal{I}(\bar{w})} (c_\ell^T \bar{w}) \right), \end{aligned} \quad (3.5)$$

where the last step follows by the construction of $\mathcal{I}(\bar{w})$. Combining (3.4) and (3.5) gives us the strict inequality

$$\max_{\ell \in \mathcal{I}(\bar{w})} (c_\ell^T w) - \max_{\ell \in L \setminus \mathcal{I}(\bar{w})} (c_\ell^T w) > 0,$$

and hence justifies the first claimed inclusion $\mathcal{I}(w) \subset \mathcal{I}(\bar{w})$.

It remains to show that $\mathcal{J}(w) \subset \mathcal{J}(\bar{w})$ for all $w \in \bar{w} + \mathcal{B}_\gamma$. The latter inclusion is obvious when $\mathcal{J}(\bar{w}) = \mathcal{J}$. Assume now that $\mathcal{J} \setminus \mathcal{J}(\bar{w}) \neq \emptyset$ and then get for every $w \in \bar{w} + \mathcal{B}_\gamma$ and $j \in \mathcal{J} \setminus \mathcal{J}(\bar{w})$ the relationships

$$w_j = \bar{w}_j + w_j - \bar{w}_j > \bar{w}_j - \gamma \geq \bar{w}_j - \min_{j \in \mathcal{J} \setminus \mathcal{J}(\bar{w})} \bar{w}_j \geq 0.$$

Thus $w_j > 0$ whenever $j \in \mathcal{J} \setminus \mathcal{J}(\bar{w})$, and thus we arrive at $\mathcal{J}(w) \subset \mathcal{J}(\bar{w})$. \triangle

The next result provides a useful representation of the condition measure (3.1) convenient for our subsequent analysis.

LEMMA 3.2. (representation of condition measure). *Assume that $\Omega \setminus S \neq \emptyset$. Then there exists $\bar{w} \in \Omega \setminus S$ such that*

$$k_F = \sup_{w \in \Omega \setminus S} \frac{\text{dist}(w; S)}{F(w)} = \frac{\text{dist}(\bar{w}; S)}{F(\bar{w})}. \quad (3.6)$$

The proof of Lemma 3.2 is based on dividing the set $\Omega \setminus S$ into a finite family of subsets and showing that the supremum is attained on each one of such sets. Before presenting this proof, we introduce some notation and prove a technical proposition.

For every $s \in S$ let $N_S(s)$ be the normal cone to S at s . Observe that this cone is polyhedral and can be represented as follows:

$$N_S(s) = \{g = w - s \mid \text{dist}(w; S) = \|w - s\|\}. \quad (3.7)$$

Also for every $s \in S$ define

$$K_s := \Omega \cap (s + N_S(s)).$$

Note that for every $s \in S$ the set K_s is a convex polytope and that

$$\Omega = \bigcup_{s \in S} K_s \quad \text{and} \quad \Omega \setminus S = \bigcup_{s \in S} (K_s \setminus \{s\}).$$

In addition, it follows from (3.7) and the definition of K_s that for every $s \in S$ and any $w \in K_s$ we have

$$\text{dist}(w; S) = \|w - s\|.$$

PROPOSITION 3.3. (supremum attainability). *For every $s \in S$ there exists $\bar{w} \in K_s \setminus \{s\}$, which realizes the supremum*

$$\sup_{w \in K_s \setminus \{s\}} \frac{\text{dist}(w; S)}{F(w)} = \frac{\text{dist}(\bar{w}; S)}{F(\bar{w})}. \quad (3.8)$$

Proof. Let $s \in S$, and let a sequence $\{w_k\} \subset K_s \setminus \{s\}$ be such that

$$\sup_{w \in K_s \setminus \{s\}} \frac{\text{dist}(w; S)}{F(w)} = \sup_{w \in K_s \setminus \{s\}} \frac{\|w - s\|}{F(w)} = \lim_{k \rightarrow \infty} \frac{\|w_k - s\|}{F(w_k)}.$$

Without loss of generality it can be assumed that $w_k \rightarrow w_0 \in K_s$ along the whole sequence $\{w_k\}$, since K_s is closed and bounded. If $w_0 \neq s$, we get $F(w_0) \neq 0$ and thus

$$\lim_{k \rightarrow \infty} \frac{\|w_k - s\|}{F(w_k)} = \frac{\|w_0 - s\|}{F(w_0)}.$$

Therefore, in this case we have found $\bar{w} = w_0$ that satisfies (3.8).

In the case when $w_0 = s$, consider the sequence $\{g_k\}$ defined by

$$g_k := \frac{w_k - s}{\|w_k - s\|} \quad \text{for all } k \in \mathbb{N}.$$

Since $\{g_k\}$ is bounded, suppose without loss of generality that $g_k \rightarrow g$ with $\|g\| = 1$. The polyhedrality of K_s implies the existence of $\lambda_1 > 0$ satisfying $s + \lambda g \in K_s$ for all $\lambda \in [0, \lambda_1]$. Further, the finiteness of the index set L allows us to get without loss of generality that

$$\mathcal{I}(w_k) = \mathcal{I}_0 \subset L \quad \text{for all } k.$$

Then there exists $l_0 \in \mathcal{I}_0$ such that

$$c_{l_0}^T w_k \geq c_{l_0}^T w_k \quad \text{whenever } l \in L \text{ and } k \in \mathbb{N}. \quad (3.9)$$

Since $w_k \rightarrow s$, we have for all $l \in L$ that

$$\lim_{k \rightarrow \infty} c_l^T w_k = c_l^T s,$$

which by (3.9) yields that $l_0 \in \mathcal{I}(s)$.

Let us show in addition that there is $\lambda_2 > 0$ for which $l_0 \in \mathcal{I}(s + \lambda g)$ whenever $\lambda \in [0, \lambda_2]$. Indeed, assume the contrary and find a sequence $\{\lambda^{(m)}\}$ with $\lambda^{(m)} \downarrow 0$ such that $l_0 \notin \mathcal{I}(s + \lambda^{(m)} g)$ for all $m \in \mathbb{N}$. Then without loss of generality (since L is finite) there exists $\bar{l} \in L$ satisfying

$$c_{\bar{l}}^T (s + \lambda^{(m)} g) > c_{l_0}^T (s + \lambda^{(m)} g) \quad \text{for all } m \in \mathbb{N}.$$

By the same arguments used above to show that $l_0 \in \mathcal{I}(s)$, we get $\bar{l} \in \mathcal{I}(s)$ and thus

$$c_{l_0}^T s = c_{\bar{l}}^T s = 0. \quad (3.10)$$

Together with (3.9), the latter yields that

$$c_{l_0}^T \frac{(w_k - s)}{\|w_k - s\|} \geq c_{\bar{l}}^T \frac{(w_k - s)}{\|w_k - s\|}.$$

By passing to the limit as $k \rightarrow \infty$, we obtain $c_{l_0}^T g \geq c_{\bar{l}}^T g$ and conclude that

$$c_{l_0}^T (s + \lambda^{(m)} g) \geq c_{\bar{l}}^T (s + \lambda^{(m)} g),$$

which contradicts our assumption and thus justifies the claim.

To proceed further, denote $\lambda := \min\{\lambda_1, \lambda_2\}$ and $\bar{w} := s + \lambda g$. Then we have

$$\frac{\text{dist}(\bar{w}; S)}{F(\bar{w})} = \frac{\|\bar{w} - s\|}{F(\bar{w})} = \frac{\lambda \|g\|}{\lambda c_{l_0}^T g} = \frac{1}{c_{l_0}^T g}.$$

But it follows at the same time that

$$\lim_{k \rightarrow \infty} \frac{\|w_k - s\|}{F(w_k)} = \lim_{k \rightarrow \infty} \frac{\|w_k - s\|}{c_{l_0}^T w_k} = \lim_{k \rightarrow \infty} \frac{1}{c_{l_0}^T \frac{w_k - s}{\|w_k - s\|}} = \frac{1}{c_{l_0}^T g},$$

which shows that \bar{w} satisfies (3.8) and thus completes the proof of the proposition. \triangle

Now let us justify our basic Lemma 3.2.

Proof of Lemma 3.2. For every $s \in S$ consider the tangent cone to Ω at s defined by

$$T_\Omega(s) := \{g \in \mathbb{R}^n \mid \exists \alpha_0 > 0 : s + \alpha g \in \Omega \text{ for all } \alpha \in (0, \alpha_0]\}.$$

We split the set S into a family of disjoint subsets \mathcal{S} , which correspond to the following equivalence classes:

$$s_1 \sim s_2 \quad \text{if and only if} \quad N_S(s_1) = N_S(s_2), \quad T_\Omega(s_1) = T_\Omega(s_2), \quad \text{and} \quad \partial F(s_1) = \partial F(s_2).$$

Observe the relationship

$$\sup_{w \in \Omega \setminus S} \frac{d(w, S)}{F(w)} = \sup_{C \in \mathcal{S}} \sup_{s \in C} \sup_{w \in K_s \setminus \{s\}} \frac{d(w, S)}{F(w)}, \quad (3.11)$$

where the outer supremum (over $C \in \mathcal{S}$) on the right-hand side is attained, since the number of different sets/classes in \mathcal{S} is finite due to the polyhedral structure of the problem. The innermost supremum in (3.11) is attained by Proposition 3.3. Hence to prove the claim, it remains to show that the supremum over $s \in C$ in (3.11) is also attained. We do it by proving that for every $C \in \mathcal{S}$ it holds

$$\sup_{w \in K_{s_1} \setminus \{s_1\}} \frac{d(w, S)}{F(w)} = \sup_{w \in K_{s_2} \setminus \{s_2\}} \frac{d(w, S)}{F(w)} \quad \text{whenever } s_1, s_2 \in C. \quad (3.12)$$

To proceed, pick an arbitrary set $C \in \mathcal{S}$ and arbitrary elements $s_1, s_2 \in C$. Fix a point $w_0 \in K_{s_1} \setminus \{s_1\}$ and get by the definition of K_{s_1} the corresponding point $g := w_0 - s_1 \in N_S(s_1)$. Since $N_S(s_1) = N_S(s_2)$ and since $N_S(s_2)$ is a cone, we have

$$\lambda g \in N_S(s_2) \text{ for all } \lambda \in [0, \infty).$$

It follows again from the definition of K_s and from the polyhedrality of Ω that $g \in T_\Omega(s_1)$ and hence $g \in T_\Omega(s_2)$ by $T_\Omega(s_1) = T_\Omega(s_2)$. This ensures the existence of $\lambda_1 > 0$ such that $s_2 + \lambda g \in \Omega$ for all $\lambda \in [0, \lambda_1]$. Thus

$$s_2 + \lambda g \in \Omega \cap (s_2 + N_S(s_2)) \setminus \{s_2\} = K_{s_2} \setminus \{s_2\} \text{ for all } \lambda \in (0, \lambda_1].$$

Since F is polyhedral, it follows that for each point s in S the function $F(\cdot) - F(s)$ is positively homogeneous in a neighborhood of s , and therefore it can be represented in this neighborhood via the subdifferential of F as

$$F(w) - F(s) = \max_{v \in \partial F(s)} v^\top (w - s).$$

Since $\partial F(s_1) = \partial F(s_2)$, find $\lambda_2 > 0$ such that $F(s_1 + \lambda g) = F(s_2 + \lambda g)$ for all $\lambda \in [0, \lambda_2]$. Letting further $\lambda := \min\{\lambda_1, \lambda_2\}$ and denoting $w_1 := s_1 + \lambda g \in K_{s_1}$ and $w_2 := s_2 + \lambda g \in K_{s_2}$, we have $F(w_1) = F(w_2)$ and

$$F(w_1) = F(s_1 + \lambda g) = F((1 - \lambda)s_1 + \lambda w_0) \leq (1 - \lambda)F(s_1) + \lambda F(w_0) = \lambda F(w_0)$$

by the convexity of F . This implies the relationships

$$\frac{\text{dist}(w_2; S)}{F(w_2)} = \frac{\text{dist}(w_1; S)}{F(w_1)} = \frac{\|\lambda g\|}{F(w_1)} \geq \frac{\lambda \|g\|}{\lambda F(w_0)} = \frac{\text{dist}(w_0; S)}{F(w_0)}.$$

Since the latter holds for any $w_0 \in K_{s_1} \setminus \{s_1\}$, we get

$$\sup_{w \in K_{s_2} \setminus \{s_2\}} \frac{\text{dist}(w; S)}{F(w)} \geq \sup_{w \in K_{s_1} \setminus \{s_1\}} \frac{\text{dist}(w; S)}{F(w)}. \quad (3.13)$$

The inverse inequality to (3.13) is obtained by interchanging the roles of s_1 and s_2 . Since our initial choice of $C \in \mathcal{S}$ and $s_1, s_2 \in C$ was arbitrary, we arrive at the equality in (3.12) for all $C \in \mathcal{S}$ and thus complete the proof of the lemma. \triangle

Now we are ready to prove our main results, namely Theorem 1.1, Theorem 1.2 and Theorem 1.3.

3.1. Proof of Theorem 1.1. We split the proof of the theorem into three major steps.

Step 1: metric regularity via condition measure. For every $\bar{w} \notin S$ and every $\bar{z} > 0$ we have the distance estimate

$$\text{dist}(\bar{w}; \Phi^{-1}(\bar{z})) \leq k_F \text{dist}(\Phi(\bar{w}); \bar{z}). \quad (3.14)$$

Proof. When $\bar{w} \notin \Omega$, the right-hand side of (3.14) becomes infinity (by the construction of Φ in (3.3) and the standard convention on $\inf \emptyset = \infty$) while the left-hand side is finite, i.e., there is nothing to prove. Considering the case of $\bar{w} \in \Omega$, observe that the left-hand side of (3.14) becomes zero when $F(\bar{w}) \leq \bar{z}$, and thus the inequality holds automatically. It remains to examine the case when $0 < \bar{z} < F(\bar{w})$ with $\bar{w} \in \Omega$.

To proceed, let $w^* := \Pi_{\Phi^{-1}(\bar{z})}(\bar{w})$, and observe that $F(w^*) = \bar{z}$, since otherwise the continuity of F would allow us to find a closer point to w in $[w^*, \bar{w}] \cap \Phi^{-1}(\bar{z})$. Thus

$$\text{dist}(\bar{w}; \Phi^{-1}(\bar{z})) = \text{dist}(\bar{w}; S(\bar{z})) = \|\bar{w} - w^*\|, \quad (3.15)$$

where $S(\cdot)$ is defined in (3.2). Let $w_0 := \Pi_S(\bar{w})$. Employing again the continuity of F , we find $\lambda \in (0, 1)$ such that

$$F(w_0 + \lambda(\bar{w} - w_0)) = \bar{z}. \quad (3.16)$$

In addition the convexity of F yields that

$$\bar{z} = F(w_0 + \lambda(\bar{w} - w_0)) \leq F(w_0) + \lambda(F(\bar{w}) - F(w_0)) = \lambda F(\bar{w}). \quad (3.17)$$

Combining the above, we have the relationships

$$\begin{aligned} \text{dist}(\bar{w}; \Phi^{-1}(\bar{z})) &= \text{dist}(\bar{w}; S(\bar{z})) \quad (\text{by (3.15)}) \\ &\leq \|w_0 + \lambda(\bar{w} - w_0) - \bar{w}\| \quad (\text{by (3.16)}) \\ &= (1 - \lambda)\|w_0 - \bar{w}\| = (1 - \lambda) \text{dist}(\bar{w}; S); \end{aligned} \quad (3.18)$$

$$\begin{aligned} \text{dist}(\Phi(\bar{w}); \bar{z}) &= F(\bar{w}) - \bar{z} \quad (\text{as } \bar{z} < F(\bar{w})) \\ &\geq F(\bar{w}) - \lambda F(\bar{w}) \quad (\text{by (3.17)}) \\ &= (1 - \lambda)F(\bar{w}), \end{aligned} \quad (3.19)$$

which finally give

$$\begin{aligned} \text{dist}(\bar{w}; \Phi^{-1}(\bar{z})) &\leq (1 - \lambda) \text{dist}(\bar{w}; S) \quad (\text{by (3.18)}) \\ &\leq (1 - \lambda)k_F F(\bar{w}) \quad (\text{as } \text{dist}(\bar{w}; S) \leq k_F F(\bar{w})) \\ &\leq k_F \text{dist}(\Phi(\bar{w}); \bar{z}) \quad (\text{by (3.19)}) \end{aligned}$$

and thus allow us to arrive at (3.14). \triangle

Step 2: distance properties. Let $\bar{w} \in \Omega \setminus S$ be such that

$$k_F = \frac{\text{dist}(\bar{w}; S)}{F(\bar{w})}, \quad (3.20)$$

let $w_0 := \Pi_S(\bar{w})$, and for $\lambda \in (0, 1)$ let

$$w_\lambda := \bar{w} + \lambda(w_0 - \bar{w}).$$

Then for any $\lambda \in (0, 1)$ we have the properties:

- (i) $F(w_\lambda) = (1 - \lambda)F(\bar{w})$.
- (ii) $\text{dist}(\bar{w}; \Phi^{-1}(F(w_\lambda))) = \lambda \text{dist}(\bar{w}; S)$.

Proof. To justify (i), observe that by the convexity of F we have

$$F(w_\lambda) \leq F(\bar{w}) + \lambda(F(w_0) - F(\bar{w})) = (1 - \lambda)F(\bar{w}) \quad (3.21)$$

in the notation above. On the other hand, the definition of k_F and the choice of \bar{w} yield

$$F(w_\lambda) \geq \frac{\text{dist}(w_\lambda; S)}{k_F} = \frac{(1 - \lambda)\|\bar{w} - w_0\|}{k_F} = \frac{(1 - \lambda) \text{dist}(\bar{w}; S)}{k_F} = (1 - \lambda)F(\bar{w}). \quad (3.22)$$

Thus assertion (i) follows from (3.21) and (3.22).

To justify (ii), it suffices to show that

$$\text{dist}(\bar{w}; \Phi^{-1}(F(w_\lambda))) = \|\bar{w} - w_\lambda\|.$$

Proceeding by contradiction, assume that $\text{dist}(\bar{w}; \Phi^{-1}(F(w_\lambda))) < \lambda \text{dist}(\bar{w}; S) = \lambda k_F F(\bar{w})$ and let $w^* := \Pi_{S(F(w_\lambda))}(\bar{w})$. By the continuity of F we have

$$\text{dist}(\bar{w}; \Phi^{-1}(F(w_\lambda))) = \text{dist}(\bar{w}; S(F(w_\lambda))) = \|\bar{w} - w^*\|,$$

which yields therefore that

$$\|\bar{w} - w^*\| = \text{dist}(\bar{w}; \Phi^{-1}(F(w_\lambda))) < \lambda k_F F(\bar{w}). \quad (3.23)$$

Taking further a point \tilde{w} closest to w^* in S , we get by (3.1) and by part (i) above that

$$\|\tilde{w} - w^*\| \leq k_F F(w^*) = k_F F(w_\lambda) = k_F(1 - \lambda)F(\bar{w}). \quad (3.24)$$

Since $\tilde{w} \in S$, the latter implies that

$$\begin{aligned} \text{dist}(\bar{w}; S) &\leq \|\bar{w} - \tilde{w}\| \\ &\leq \|\bar{w} - w^*\| + \|w^* - \tilde{w}\| \quad (\text{by the triangle inequality}) \\ &< \lambda k_F F(\bar{w}) + (1 - \lambda)k_F F(\bar{w}) \quad (\text{by (3.23) and (3.24)}) \\ &= k_F F(\bar{w}), \end{aligned}$$

which contradicts (3.20) and thus completes the proof of Step 2. \triangle

Step 3: condition measure via metric regularity. *We have the equality*

$$k_F = \sup_{w \in \Omega \setminus S} \text{reg } \Phi(w, F(w)). \quad (3.25)$$

Proof. Let us first show that

$$k_F \geq \sup_{w \in \Omega \setminus S} \text{reg } \Phi(w, F(w)). \quad (3.26)$$

Assuming the contrary, find $(w', z') \in \text{gph } F$, $w' \in \Omega \setminus S$ satisfying

$$\text{reg } \Phi(w', z') > k_F.$$

Observe that there exists a neighborhood of (w', z') such that for all points (w, z) in that neighborhood $w \notin S$ and $z > 0$. By the definition of metric regularity we can find \bar{w}, \bar{z} in such a neighborhood of (w', z') for which

$$\text{dist}(\bar{w}, \Phi^{-1}(\bar{z})) > k_F \text{dist}(\Phi(\bar{w}); \bar{z}).$$

The latter contradicts Step 1 and thus ensures (3.26).

To prove the opposite inequality in (3.25), by Lemma 3.2 find $\bar{w} \in \Omega \setminus S$ such that

$$\text{dist}(\bar{w}; S) = k_F F(\bar{w}).$$

Let $w_0 := \Pi_S(\bar{w})$ and define

$$w_\lambda := \bar{w} + \lambda(w_0 - \bar{w}), \quad 0 < \lambda < 1.$$

It follows from Step 2 that for every $\lambda \in (0, 1)$ and the above choice of \bar{w} we have

$$\frac{\text{dist}(\bar{w}; \Phi^{-1}(F(w_\lambda)))}{\text{dist}(F(w_\lambda); \Phi(\bar{w}))} = \frac{\lambda \text{dist}(\bar{w}; S)}{\lambda F(\bar{w})} = k_F.$$

The latter implies, since $w_\lambda \rightarrow \bar{w}$ and $F(w_\lambda) \rightarrow F(\bar{w})$ as $\lambda \downarrow 0$, that

$$\text{reg } \Phi(\bar{w}, F(\bar{w})) \geq \limsup_{\lambda \downarrow 0} \frac{\text{dist}(\bar{w}; \Phi^{-1}(F(w_\lambda)))}{\text{dist}(F(w_\lambda); \Phi(\bar{w}))} = k_F,$$

which therefore yields

$$k_F \leq \sup_{w \in \Omega \setminus S} \text{reg } \Phi(w, F(w))$$

and completes the the proof of the theorem. \triangle

3.2. Proof of Theorem 1.2. First of all, observe by Step 1 in the proof of Theorem 1.1 that the multifunction Φ is metrically regular around $(w, z) \in \text{gph } \Phi$ for every $w \in \Omega \setminus S$. Employing the corresponding results of Section 2, for $(w, z) \in \text{gph } \Phi$ with $w \in \Omega \setminus S$ we get

$$\begin{aligned} \text{reg } \Phi(w, z) &= \text{lip } \Phi^{-1}(z, w) \quad (\text{by Theorem 2.2}) \\ &= \|D^*\Phi^{-1}(z, w)\| \quad (\text{by Theorem 2.6}) \\ &= \|(D^*\Phi(w, z))^{-1}\| \quad (\text{by the definition of } D^*\Phi(w, z)) \\ &= \sup_{|v|=1} \frac{1}{\text{dist}(0; D^*\Phi(w, z)(v))} \quad (\text{by Proposition 2.5 with } n = 1). \end{aligned}$$

This gives therefore the regularity exact bound formula

$$\text{reg } \Phi(w, z) = \frac{1}{\min \{ \text{dist}(0; D^*\Phi(w, z)(-1)), \text{dist}(0; D^*\Phi(w, z)(1)) \}}. \quad (3.27)$$

Defined next a set-valued mapping $\tilde{F}: \mathbb{R}^{m+n} \rightrightarrows \mathbb{R}$ by

$$\tilde{F}(w) := [F(w), \infty) \quad \text{with } \text{gph } \tilde{F} = \text{epi } F \quad (3.28)$$

and observe that it is Lipschitz-like at every point of its graph, which is the epigraph of a Lipschitz continuous function. Furthermore, the graph of \tilde{F} is convex, and hence \tilde{F} is graphically regular at any point of its graph by Proposition 2.3, which also ensures the normal regularity of the convex set Ω . Applying Proposition 2.4 to the sum $\Phi = \tilde{F} + \delta_\Omega$, we get the equality

$$D^*\Phi(w, z)(\lambda) = D^*\tilde{F}(w, z)(\lambda) + N_\Omega(w) \quad \text{for all } \lambda \in \mathbb{R}. \quad (3.29)$$

It follows from the structure of \tilde{F} in (3.28), the coderivative definition (2.4), and the well-known subdifferential representation

$$\partial\varphi(\bar{x}) = \{v \in \mathbb{R}^n \mid (v, -1) \in N((\bar{x}, \varphi(\bar{x})); \text{epi } \varphi)\}, \quad \bar{x} \in \text{dom } \varphi,$$

for any convex function $\varphi: \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$ that

$$D^*\tilde{F}(w, z)(1) = \begin{cases} \partial F(w), & z = F(w), \\ \emptyset, & z > F(w), \end{cases} \quad D^*\tilde{F}(w, z)(-1) = \emptyset. \quad (3.30)$$

Combining (3.27), (3.29), and (3.30) gives us the formula

$$\text{reg } \Phi(w, z) = \frac{1}{\text{dist}(0; \partial F(w) + N_\Omega(w))}, \quad (3.31)$$

which is (1.10). \triangle

3.3. Proof of Theorem 1.3.. By Theorem 1.1 and Theorem 1.2, it suffices to show the following representations for $\partial F(x, y)$ and $N_{\Delta_m \times \Delta_n}(x, y)$:

$$\partial F(x, y) = \text{co} \{ (a_i, b_k) \in \mathbb{R}^m \times \mathbb{R}^n \mid i \in I(x), k \in K(y) \}, \quad (3.32)$$

$$N_{\Delta_m \times \Delta_n}(x, y) = \text{span} \{ \mathbf{1}_m \} \times \text{span} \{ \mathbf{1}_n \} - \text{cone} [\text{co} \{ e_j \mid j \in J(x, y) \}]. \quad (3.33)$$

Indeed, the classical subdifferential formula for max-functions (see, e.g., [19, Exercise 8.31]) gives us

$$\partial \left(\max_{\ell \in L} c_\ell^\text{T} w \right) = \text{co} \{ c_{\bar{\ell}} \mid \bar{\ell} \in L, c_{\bar{\ell}}^\text{T} w = \max_{\ell \in L} c_\ell^\text{T} w \}.$$

This implies by the max-structure of the function F in (1.4) that

$$\partial F(x, y) = \text{co} \{ (a_i, b_k) \mid i \in I(x), k \in K(y) \},$$

which is (3.32).

To prove (3.33), we recall first the calculus formula

$$N_{A \cap B}(w) = N_A(w) + N_B(w)$$

held at every $w \in A \cap B$ for arbitrary convex polyhedra in finite dimensions; see, e.g., [18, Corollary 23.8.1]). Thus we have in our case that

$$N_\Omega(w) = N_{\{Eu=f\}}(w) + N_{\{u \geq 0\}}(w).$$

Moreover, it is easy to see that

$$N_{\{Eu=f\}}(w) = (\ker E)^\perp = \text{span} \{ \mathbf{1}_m \} \times \text{span} \{ \mathbf{1}_n \},$$

$$N_{\{u \geq 0\}}(w) = -\text{cone} [\text{co} \{ e_j \mid j \in \mathcal{J}(w) \}].$$

Thus for any $w = (x, y) \in \Omega = \Delta_n \times \Delta_m$ we have

$$N_{\Delta_m \times \Delta_n}(x, y) = \text{span} \{ \mathbf{1}_m \} \times \text{span} \{ \mathbf{1}_n \} - \text{cone} [\text{co} \{ e_j \mid j \in \mathcal{J}(x, y) \}],$$

which is (3.33). \triangle

4. Condition measure formula via alternative proof. In this section we give another proof of Theorem 1.3 based on convex optimization. This proof is split into three lemmas and the preceding technical claim.

Given a point $\bar{w} \in \Omega \setminus S$ and keeping the notation above, consider the following two problems of parametric optimization (with the parameter $z \in \mathbb{R}$) defined by

$$\begin{array}{ll} V_z(\bar{w}) := \min_w \|w - \bar{w}\| & \tilde{V}_z(\bar{w}) := \min_w \|w - \bar{w}\| \\ \text{s.t. } c_\ell^\top w \leq z \quad \forall \ell \in \mathcal{I}(\bar{w}) & \text{s.t. } \max_{\ell \in L} \{ c_\ell^\top w \} = z \\ Ew = f & Ew = f \\ w_j \geq 0 \quad \forall j \in \mathcal{J}(\bar{w}) & w \geq 0 \end{array} \quad (P_z) \quad \text{and} \quad (\tilde{P}_z)$$

and name (P_z) and (\tilde{P}_z) the *first* and *second parametric problem*, respectively. Observe that for every $\bar{w} \in \Omega \setminus S$ and $z \in \mathbb{R}_+$ the optimal value $\tilde{V}_z(\bar{w})$ in problem (\tilde{P}_z) is equal to $\text{dist}(\bar{w}; S(z))$. Although the proof of the following claim is straightforward, we provide it for completeness and the reader's convenience.

CLAIM 4.1. (stability of optimal solutions to first parametric problem). *For any $\bar{w} \in \Omega \setminus S$ and any $\gamma > 0$ there is $\varepsilon > 0$ (depending on \bar{w} and γ) such that whenever $z \in [F(\bar{w}) - \varepsilon, F(\bar{w})]$ a unique solution w_z to problem (P_z) exists and satisfies the continuity property $\|w_z - \bar{w}\| \leq \gamma$ with respect to the parameter z .*

Proof. Fix $\gamma > 0$ and put $w^S := \Pi_S(\bar{w})$. Let

$$\tilde{w} := \bar{w} + \tau(w^S - \bar{w}) \quad \text{with} \quad \tau := \min \left\{ \frac{\gamma}{\|w^S - \bar{w}\|}, 1 \right\} \in (0, 1].$$

Setting $\varepsilon := F(\bar{w}) - F(\tilde{w})$, observe by the convexity of F that

$$\varepsilon = F(\bar{w}) - F(\tilde{w}) \geq \tau F(\bar{w}) > 0.$$

We have furthermore that

$$E\tilde{w} = E\bar{w} + \tau(Ew^S - E\bar{w}) = f,$$

$$\tilde{w}_j = (1 - \tau)\bar{w}_j + \tau w_j^S \geq 0 \text{ for all } j \in \mathcal{J}(\bar{w}), \text{ and}$$

$$\|\tilde{w} - \bar{w}\| = \tau\|w^S - \bar{w}\| \leq \gamma,$$

which imply the inclusion

$$\tilde{w} \in \Delta := \{w \mid \|w - \bar{w}\| \leq \gamma, \quad Ew = f, \quad w_j \geq 0 \text{ for all } j \in \mathcal{J}(\bar{w})\}.$$

Since the set Δ is obviously convex with $\bar{w} \in \Delta$, we get

$$w_t := \bar{w} + t(\tilde{w} - \bar{w}) \in \Delta \text{ whenever } t \in [0, 1].$$

It follows from $F(\tilde{w}) = F(\bar{w}) - \varepsilon$ and the continuity of F that for every $z \in [F(\bar{w}) - \varepsilon, F(\bar{w})]$ there is $t_z \in [0, 1]$ such that w_{t_z} satisfies the equation

$$F(w_{t_z}) = z.$$

For any z from the above we easily get that w_{t_z} is feasible to problem (P_z) , that the set of feasible solutions to this problem is surely closed and bounded, and that the cost function is continuous with respect to w . Thus (P_z) admits an optimal solution, which is unique as a unique projection of \bar{w} on the convex feasible set. Finally,

$$V_z(\bar{w}) = \|w_z - \bar{w}\| \leq \|w_{t_z} - \bar{w}\| \leq t_z \|\tilde{w} - \bar{w}\| \leq \gamma,$$

and hence the optimal solution w_z belongs to the ball $\bar{w} + \mathcal{B}_\gamma$. \triangle

The next result, whose proof is based on Claims 3.1 and 4.1, indicates the parameter region on which the optimal values in the first and second parametric problems agree.

LEMMA 4.2. (optimal values agree for both parametric problems). *Let $\bar{w} \in \Omega \setminus S$. Then there exists $\varepsilon_{\bar{w}} \in (0, F(\bar{w}))$ such that for every parameter $z \in [F(\bar{w}) - \varepsilon_{\bar{w}}, F(\bar{w})]$ the optimal values of problems (P_z) and (\tilde{P}_z) coincide.*

Proof. Fix $\bar{w} \in \Omega \setminus S$ and observe that the set of feasible solutions for (\tilde{P}_z) obviously belongs to the set of feasible solutions for (P_x) . Thus we have $\tilde{V}_z(\bar{w}) \geq V_z(\bar{w})$ for all $z \in \mathbb{R}$. It remains to show that there exists $\varepsilon_{\bar{w}} > 0$ such that $\tilde{V}_z(\bar{w}) \leq V_z(\bar{w})$ whenever $z \in [F(\bar{w}) - \varepsilon_{\bar{w}}, F(\bar{w})]$.

Employing Claim 3.1, find $\gamma > 0$ for which $\mathcal{I}(w) \subset \mathcal{I}(\bar{w})$ and $\mathcal{J}(w) \subset \mathcal{J}(\bar{w})$ when $w \in \bar{w} + \mathcal{B}_\gamma$. Further, it follows from Claim 4.1 that for such γ there is $\varepsilon > 0$ with the property: whenever $z \in [F(\bar{w}) - \varepsilon, F(\bar{w})]$ there exists a unique solution w_z to problem (P_z) satisfying $w_z \in \bar{w} + \mathcal{B}_\gamma$. Our choice of γ ensures the feasibility of w_z in problem (\tilde{P}_z) , and therefore we have the relationships

$$\tilde{V}_z(\bar{w}) \leq \|w_z - \bar{w}\| = V_z(\bar{w}),$$

which thus complete the proof of the lemma. \triangle

LEMMA 4.3. (distances to solution sets). *For every $w \in \Omega \setminus S$ denote $z_w := F(w) - \varepsilon_w$ with ε_w taken from Lemma 4.2. Then we have*

$$\sup_{w \in \Omega \setminus S} \frac{\text{dist}(w; S(z_w))}{F(w) - z_w} = \sup_{w \in \Omega \setminus S} \frac{\text{dist}(w; S)}{F(w)}. \quad (4.1)$$

Proof. Fix $w \in \Omega \setminus S$ and $z \in (0, F(w))$, and then let

$$k(w, z) := \frac{\text{dist}(w; S(z))}{F(w) - z}.$$

We first justify the inequality

$$\sup_{w \in \Omega \setminus S} k(w, z_w) \leq \sup_{w \in \Omega \setminus S} k(w, 0), \quad (4.2)$$

which gives the one in (4.1). Pick any $\bar{w} \in \Omega \setminus S$, let $w^S := \arg \min_{w \in S} \|w - \bar{w}\|$ and

$$w_t := (1 - t)w^S + t\bar{w}, \quad t \in [0, 1].$$

Since $z_{\bar{w}} \in (0, F(\bar{w}))$, there is $\tau \in (0, 1)$ with $F(w_\tau) = z_{\bar{w}}$. By the convexity of F we have

$$z_{\bar{w}} = F(w_\tau) \leq (1 - \tau)F(w^S) + \tau F(\bar{w}) = \tau F(\bar{w}). \quad (4.3)$$

Further, it follows from $w_\tau \in S(z_{\bar{w}})$ that

$$\text{dist}(\bar{w}; S(z_{\bar{w}})) \leq \|\bar{w} - w_\tau\| = (1 - \tau)\|w^S - \bar{w}\| = (1 - \tau) \text{dist}(\bar{w}; S). \quad (4.4)$$

Combining (4.3) and (4.4) gives us

$$k(\bar{w}, z_{\bar{w}}) = \frac{\text{dist}(\bar{w}; S(z_{\bar{w}}))}{\text{dist}(F(\bar{w}); z_{\bar{w}})} \leq \frac{(1 - \tau) \text{dist}(\bar{w}; S)}{(1 - \tau)F(\bar{w})} = k(\bar{w}, 0),$$

which yields (4.2) and thus the corresponding inequality in (4.1).

It remains to show that

$$\sup_{w \in \Omega \setminus S} k(w, z_w) \geq \sup_{w \in \Omega \setminus S} k(w, 0), \quad (4.5)$$

which ensures the equality in (4.1). By Lemma 3.2 we have that $k_F = \sup_{w \in \Omega \setminus S} k(w, 0)$ and that the maximum is attained at some $\bar{w} \in \Omega \setminus S$. Given $z \in (0, F(\bar{w}))$, let

$$w_z := \arg \min_{w \in S(z_w)} \|w - \bar{w}\|$$

and observe the estimate

$$\text{dist}(\bar{w}; S) \leq \|\bar{w} - w_z\| + \text{dist}(w_z; S),$$

implying in turn that

$$\sup_{w \in \Omega \setminus S} k(w, z_w) \geq k(\bar{w}, z_{\bar{w}}) = \frac{\|\bar{w} - w_{z_{\bar{w}}}\|}{F(\bar{w}) - z_{\bar{w}}} \geq \frac{\text{dist}(\bar{w}; S) - \text{dist}(w_{z_{\bar{w}}}; S)}{F(\bar{w}) - z_{\bar{w}}}.$$

On the other hand, we have the equality $\text{dist}(\bar{w}; S) = F(\bar{w})k(\bar{w}, 0)$ by the definition of $k(w, z)$ and also the relationships

$$\text{dist}(w_{z_{\bar{w}}}; S) = F(w_{z_{\bar{w}}})k(w_{z_{\bar{w}}}, 0) \leq z_{\bar{w}}k(\bar{w}, 0)$$

due to $F(w_{z_{\bar{w}}}) = z_{\bar{w}}$ and $k(\bar{w}, 0) = \sup_{w \in \Omega \setminus S} k(w, 0)$. Thus

$$\sup_{w \in \Omega \setminus S} k(w, z_w) \geq \frac{F(\bar{w})k(\bar{w}, 0) - z_{\bar{w}}k(\bar{w}, 0)}{F(\bar{w}) - z_{\bar{w}}} = k(\bar{w}, 0) = \sup_{w \in \Omega \setminus S} k(w, 0),$$

which justifies (4.5) and completes the proof of the lemma. \triangle

The last lemma establishes, by employing Lagrangian duality, a precise formula for computing the optimal value of the cost function in the parametric problem (P_z) —and hence in (\tilde{P}_z) —via the initial data.

LEMMA 4.4. (computing optimal values of parametric problems). *Let $\bar{w} \in \Omega \setminus S$ and $z \in (0, F(\bar{w}))$. Then the optimal value $V_z(\bar{w})$ of problem (P_z) is computed by*

$$V_z(\bar{w}) = \frac{F(\bar{w}) - z}{\text{dist}(0; \text{co}\{c_i, i \in \mathcal{I}(\bar{w})\} + (\ker E)^\perp - \text{cone}[\text{co}\{e_j, j \in \mathcal{J}(\bar{w})\}])}.$$

Proof. Observe that problem (P_z) can be reformulated as

$$V_z(\bar{w}) = \inf_w \sup_{\substack{\|u\| \leq 1, \\ \lambda_i \geq 0, i \in \mathcal{I}(\bar{w}), \\ v \in \mathbb{R}^m, \\ \mu_j \geq 0, j \in \mathcal{J}(\bar{w}), \\ \mu_j = 0, j \in \mathcal{J} \setminus \mathcal{J}(\bar{w})}} \left[u^\top(w - \bar{w}) + \sum_{i \in \mathcal{I}(\bar{w})} \lambda_i (c_i^\top w - z) + v^\top(Ew - f) - \sum_{j \in \mathcal{J}(\bar{w})} \mu_j w_j \right].$$

Observe that the convex optimization problem (P_z) satisfies the Slater condition: For $\delta \in (0, 1)$ sufficiently small and $\tilde{w} \in S$, the point $(1-\delta)\tilde{w} + \delta(\frac{1}{n}\mathbf{1}_n, \frac{1}{m}\mathbf{1}_m)$ is a strictly feasible point. Therefore, we can interchange the supremum and the infimum above by Lagrangian duality. This gives

$$V_z(\bar{w}) = \sup_{\substack{\|u\| \leq 1, \\ \lambda_i \geq 0, i \in \mathcal{I}(\bar{w}), \\ v \in \mathbb{R}^m, \\ \mu_j \geq 0, j \in \mathcal{J}(\bar{w}), \\ \mu_j = 0, j \in \mathcal{J} \setminus \mathcal{J}(\bar{w})}} \inf_w \left[u^\top(w - \bar{w}) + \sum_{i \in \mathcal{I}(\bar{w})} \lambda_i (c_i^\top w - z) + v^\top(Ew - f) - \sum_{j \in \mathcal{J}(\bar{w})} \mu_j w_j \right].$$

Regrouping the terms inside the square brackets, we obtain

$$V_z(\bar{w}) = \sup_{\substack{\|u\| \leq 1, \\ \lambda_i \geq 0, i \in \mathcal{I}(\bar{w}), \\ v \in \mathbb{R}^m, \\ \mu_j \geq 0, j \in \mathcal{J}(\bar{w}), \\ \mu_j = 0, j \in \mathcal{J} \setminus \mathcal{J}(\bar{w})}} \inf_w \left[\left(u + \sum_{i \in \mathcal{I}(\bar{w})} \lambda_i c_i + E^\top v - \mu \right)^\top w - u^\top \bar{w} - \sum_{i \in \mathcal{I}(\bar{w})} \lambda_i z - v^\top f \right].$$

Observe further that, whenever the term $(u + \sum_{i \in \mathcal{I}(\bar{w})} \lambda_i c_i + E^\top v - \mu)$ is not zero, the inner infimum in w necessarily becomes $-\infty$. This allows us to put

$$u = - \sum_{i \in \mathcal{I}(\bar{w})} \lambda_i c_i - E^\top v + \mu$$

and consequently rewrite the expression for $V_z(\bar{w})$ as follows:

$$V_z(\bar{w}) = \sup_{\substack{\|\sum_{i \in \mathcal{I}(\bar{w})} \lambda_i c_i + E^\top v - \mu\| \leq 1, \\ \lambda_i \geq 0, i \in \mathcal{I}(\bar{w}), \\ \mu_j \geq 0, j \in \mathcal{J}(\bar{w}), \\ \mu_j = 0, j \in \mathcal{J} \setminus \mathcal{J}(\bar{w})}} \left[\left(\sum_{i \in \mathcal{I}(\bar{w})} \lambda_i c_i + E^\top v - \mu \right)^\top \bar{w} - \sum_{i \in \mathcal{I}(\bar{w})} \lambda_i z - v^\top f \right].$$

Regrouping again gives us the formula

$$V_z(\bar{w}) = \sup_{\substack{\|\sum_{i \in \mathcal{I}(\bar{w})} \lambda_i c_i + E^T v - \mu\| \leq 1, \\ \lambda_i \geq 0, i \in \mathcal{I}(\bar{w}), \\ \mu_j \geq 0, j \in \mathcal{J}(\bar{w}), \\ \mu_j = 0, j \in \mathcal{J} \setminus \mathcal{J}(\bar{w})}} \left[\sum_{i \in \mathcal{I}(\bar{w})} \lambda_i (c_i^T \bar{w} - z) + v^T (E \bar{w} - f) - \mu^T \bar{w} \right].$$

Noting that $E \bar{w} = f$, $\mu^T \bar{w} = 0$ and $c_i^T \bar{w} = F(\bar{w})$ for $i \in \mathcal{I}(\bar{w})$, we have

$$V_z(\bar{w}) = (F(\bar{w}) - z) \sup_{\substack{\|\sum_{i \in \mathcal{I}(\bar{w})} \lambda_i c_i + E^T v - \mu\| \leq 1, \\ \lambda_i \geq 0, i \in \mathcal{I}(\bar{w}), \\ \mu_j \geq 0, j \in \mathcal{J}(\bar{w}), \\ \mu_j = 0, j \in \mathcal{J} \setminus \mathcal{J}(\bar{w})}} \sum_{i \in \mathcal{I}(\bar{w})} \lambda_i.$$

Since $V_z(\bar{w}) \neq 0$, the latter yields

$$V_z(\bar{w}) = (F(\bar{w}) - z) \sup_{\substack{\frac{\sum_{i \in \mathcal{I}(\bar{w})} \lambda_i c_i}{\sum_{i \in \mathcal{I}(\bar{w})} \lambda_i} + E^T \frac{v}{\sum_{i \in \mathcal{I}(\bar{w})} \lambda_i} - \frac{\mu}{\sum_{i \in \mathcal{I}(\bar{w})} \lambda_i} \leq \frac{1}{\sum_{i \in \mathcal{I}(\bar{w})} \lambda_i}, \\ \lambda_i \geq 0, i \in \mathcal{I}(\bar{w}), \\ \mu_j \geq 0, j \in \mathcal{J}(\bar{w}), \\ \mu_j = 0, j \in \mathcal{J} \setminus \mathcal{J}(\bar{w})}} \frac{1}{1 / \sum_{i \in \mathcal{I}(\bar{w})} \lambda_i},$$

which can be written as

$$V_z(\bar{w}) = (F(\bar{w}) - z) \sup_{\substack{\lambda_i \geq 0, i \in \mathcal{I}(\bar{w}), \\ \mu_j \geq 0, j \in \mathcal{J}(\bar{w}), \\ \mu_j = 0, j \in \mathcal{J} \setminus \mathcal{J}(\bar{w})}} \frac{1}{\left\| \frac{\sum_{i \in \mathcal{I}(\bar{w})} \lambda_i c_i}{\sum_{i \in \mathcal{I}(\bar{w})} \lambda_i} + E^T \frac{v}{\sum_{i \in \mathcal{I}(\bar{w})} \lambda_i} - \frac{\mu}{\sum_{i \in \mathcal{I}(\bar{w})} \lambda_i} \right\|}.$$

Changing further the variables by

$$\tilde{\lambda}_i := \frac{\lambda_i}{\sum_{i \in \mathcal{I}(\bar{w})} \lambda_i}, \quad \tilde{\mu} := \frac{\mu}{\sum_{i \in \mathcal{I}(\bar{w})} \lambda_i}, \quad \tilde{v} := \frac{v}{\sum_{i \in \mathcal{I}(\bar{w})} \lambda_i},$$

we arrive at the expression

$$V_z(\bar{w}) = (F(\bar{w}) - z) \sup_{\substack{\tilde{\lambda}_i \geq 0, i \in \mathcal{I}(\bar{w}), \\ \sum_{i \in \mathcal{I}(\bar{w})} \tilde{\lambda}_i = 1, \\ \tilde{\mu}_j \geq 0, j \in \mathcal{J}(\bar{w}), \\ \tilde{\mu}_j = 0, j \in \mathcal{J} \setminus \mathcal{J}(\bar{w})}} \frac{1}{\left\| \sum_{i \in \mathcal{I}(\bar{w})} \tilde{\lambda}_i c_i + E^T \tilde{v} - \tilde{\mu} \right\|},$$

which can be equivalently written as

$$V_z(\bar{w}) = \frac{F(\bar{w}) - z}{\inf_{\substack{\tilde{\lambda}_i \geq 0, i \in \mathcal{I}(\bar{w}), \\ \sum_{i \in \mathcal{I}(\bar{w})} \tilde{\lambda}_i = 1, \\ \tilde{\mu}_j \geq 0, j \in \mathcal{J}(\bar{w}), \\ \tilde{\mu}_j = 0, j \in \mathcal{J} \setminus \mathcal{J}(\bar{w})}} \left\| \sum_{i \in \mathcal{I}(\bar{w})} \tilde{\lambda}_i c_i + E^T \tilde{v} - \tilde{\mu} \right\|}.$$

Recalling the notation of Section 1 allows us to reduce the latter expression to the one in the lemma formulation and thus finish the proof. \triangle

Combining the obtained lemmas with the definitions above, we can now complete the alternative proof of the condition measure formula in Theorem 1.3.

Proof of Theorem 1.3. For every $w \in \Omega \setminus S$ choose the parameter z_w as in Lemma 4.3, i.e., put $z_w = F(w) - \varepsilon_w$, where ε_w is taken from Lemma 4.2. Then we have

$$\begin{aligned}
k_F &= \inf \{k \geq 0 \mid \text{dist}(w; S) \leq kF(w) \text{ for all } w \in \Omega \setminus S\} \quad (\text{by definition (3.1)}) \\
&= \sup_{w \in \Omega \setminus S} \frac{\text{dist}(w; S)}{F(w)} \quad (\text{as } \Omega \setminus S \neq \emptyset) \\
&= \sup_{w \in \Omega \setminus S} \frac{\text{dist}(w; S(z_w))}{F(w) - z_w} \quad (\text{by Lemma 4.3}) \\
&= \sup_{w \in \Omega \setminus S} \frac{\tilde{V}_{z_w}(w)}{F(w) - z_w} \quad (\text{by the definition of } (P'_z)) \\
&= \sup_{w \in \Omega \setminus S} \frac{V_{z_w}(w)}{F(w) - z_w} \quad (\text{by Lemma 4.2 and the choice of } z_w) \\
&= \sup_{w \in \Omega \setminus S} \frac{1}{\text{dist}(0; \text{co}_{i \in \mathcal{I}(w)}\{c_\ell\} + (\ker E)^\perp - \text{cone}[\text{co}_{j \in \mathcal{J}(w)}\{e_j\}])} \quad (\text{by Lemma 4.4}).
\end{aligned}$$

Letting $w = (x, y) \in \Omega$, observe finally that

$$\{c_\ell \mid \ell \in \mathcal{I}(w)\} = \{(a_i, b_k) \mid i \in I(x), k \in K(y)\}, \quad (\ker E)^\perp = \text{span}\{\mathbf{1}_n\} \times \text{span}\{\mathbf{1}_m\},$$

and $\kappa(A) = k_F$, which complete the proof of the theorem. \triangle

Acknowledgments. The authors are indebted to two anonymous referees for their valuable suggestions and remarks that allowed us to essentially improve the original presentation. We also gratefully acknowledge helpful discussions with Arkadi Nemirovski and Yurii Nesterov on the topics and results of this paper.

REFERENCES

- [1] F. J. ARAGÓN ARTACHO, A. L. DONTCHEV AND M. H. GEOFFROY, *Convergence of the proximal point method for metrically regular mappings*, ESAIM Proc., 17 (2007), pp. 1–8.
- [2] H. ATTOUCH, J. BOLTE, P. REDONT AND A. SOUBEYRAN, *Alternating minimization and projection methods for nonconvex problems, An approach based on the Kurdyka-Lojasiewicz inequality*, Math. Oper. Res., to appear.
- [3] A. L. DONTCHEV, A. D. LEWIS AND R. T. ROCKAFELLAR, *The radius of metric regularity*. Trans. Amer. Math. Soc., 355 (2003), pp. 493–517.
- [4] A. GILPIN, J. PEÑA AND T. SANDHOLM, *First-order algorithm with $O(\ln(1/\varepsilon))$ convergence for ε -equilibrium in two-person zero-sum games*, in Proc. 23rd Nat. Conf. Art. Intel. (AAAI), 2008, pp. 75–82.
- [5] N. HIGHAM, *Accuracy and Stability of Numerical Algorithms*, SIAM, Philadelphia, PA, 1996.
- [6] D. KLATTE AND B. KUMMER, *Optimization methods and stability of inclusions in Banach spaces*, Math. Program., 117 (2009), pp. 305–330.
- [7] A. S. LEWIS, D. R. LUKE AND J. MALICK, *Local linear convergence for alternating and averaged nonconvex projections*, Found. Comp. Math., 9 (2009), pp. 485–513.
- [8] B. S. MORDUKHOVICH, *Maximum principle in the problem of time optimal response with nonsmooth constraints*, Appl. Math. Mech., 40 (1976), 960–969.
- [9] B. S. MORDUKHOVICH, *Complete characterizations of openness, metric regularity, and Lipschitzian properties of multifunctions*, Trans. Amer. Math. Soc., 340 (1993), pp. 1–35.
- [10] B. S. MORDUKHOVICH, *Variational Analysis and Generalized Differentiation, I: Basic Theory*, Grundlehren Series (Fundamental Principles of Mathematical Sciences), Vol. 330, Springer, Berlin, 2006.
- [11] YU. NESTEROV, *Excessive gap technique in nonsmooth convex minimization*, SIAM J. Optim., 16 (2005), pp. 235–249.

- [12] YU. NESTEROV, *Smooth minimization of nonsmooth functions*, Math. Program., 103 (2005), pp. 127–152.
- [13] M. OSBORNE AND A. RUBINSTEIN. *A Course in Game Theory*. MIT Press, Cambridge, MA, 1994.
- [14] J. PEÑA AND J. RENEGAR, *Computing approximate solutions for convex conic systems of constraints*, Math. Program., 87 (2000), pp. 351–383.
- [15] J. RENEGAR, *Incorporating condition measures into the complexity theory of linear programming*, SIAM J. Optim., 5 (1995), pp. 506–524.
- [16] J. RENEGAR, *Linear programming, complexity theory and elementary functional analysis*, Math. Program., 70 (1995), pp. 279–351.
- [17] S. M. ROBINSON, *Strongly regular generalized equations*, Math. Oper. Res., 5 (1980), pp. 43–62.
- [18] R. T. ROCKAFELLAR, *Convex Analysis*, Princeton University Press, Princeton, NJ, 1970.
- [19] R. T. ROCKAFELLAR AND R. J-B. WETS, *Variational Analysis*, Grundlehren Series (Fundamental Principles of Mathematical Sciences), Vol. 317, Springer, Berlin, 1998.
- [20] I. ROMANOVSKII, *Reduction of a game with complete memory to a matrix game*, Soviet Mathematics, 3 (1962), pp. 678–681.
- [21] M. SHUB AND S. SMALE, *Complexity of Bézout's theorem, I: Geometric aspects*, J. Amer. Math. Soc., 6 (1993), pp. 459–501.
- [22] B. VON STENGEL, *Efficient computation of behavior strategies*, Games and Economic Behavior, 14 (2006), pp. 220–246.
- [23] B. VON STENGEL, *Equilibrium computation for two-player games in strategic and extensive form*, in Algorithmic Game Theory, N. Nisan et al., eds, Cambridge University Press, Cambridge, UK, 2007, pp. 53–78.